



Capturing Biomarkers and Molecular Targets in Cellular Landscapes From Dynamic Reaction Network Models and Machine Learning

Susan D. Mertins^{1,2,3*}†

¹ Department of Science, Mount St. Mary's University, Emmitsburg, MD, United States, ² Biomedical Informatics and Data Science Directorate, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Limited Liability Company (LLC), Frederick, MD, United States, ³ BioSystems Strategies, Limited Liability Company (LLC), Frederick, MD, United States

OPEN ACCESS

Edited by:

Mónica Hebe Vazquez-Levin,
Consejo Nacional de Investigaciones
Científicas y Técnicas (CONICET),
Argentina

Reviewed by:

Mohit Kumar Jolly,
Indian Institute of Science (IISc), India

*Correspondence:

Susan D. Mertins
smertins@biosystemsstrategies.com

†Present address:

Susan D. Mertins,
BioSystems Strategies, LLC,
Frederick, MD, United States

Specialty section:

This article was submitted to
Cancer Genetics,
a section of the journal
Frontiers in Oncology

Received: 30 October 2021

Accepted: 31 December 2021

Published: 21 January 2022

Citation:

Mertins SD (2022) Capturing
Biomarkers and Molecular
Targets in Cellular Landscapes
From Dynamic Reaction Network
Models and Machine Learning.
Front. Oncol. 11:805592.
doi: 10.3389/fonc.2021.805592

Computational dynamic ODE models of cell function describing biochemical reactions have been created for decades, but on a small scale. Still, they have been highly effective in describing and predicting behaviors. For example, oscillatory phospho-ERK levels were predicted and confirmed in MAPK signaling encompassing both positive and negative feedback loops. These models typically were limited and not adapted to large datasets so commonly found today. But importantly, ODE models describe reaction networks in well-mixed systems representing the cell and can be simulated with ordinary differential equations that are solved deterministically. Stochastic solutions, which can account for noisy reaction networks, in some cases, also improve predictions. Today, dynamic ODE models rarely encompass an entire cell even though it might be expected that an upload of the large genomic, transcriptomic, and proteomic datasets may allow whole cell models. It is proposed here to combine output from simulated dynamic ODE models, completed with omics data, to discover both biomarkers in cancer *a priori* and molecular targets in the Machine Learning setting.

Keywords: biomarkers, molecular targets, drug discovery, drug development, pharmacodynamic modeling, ODE modeling, machine learning

INTRODUCTION

Understanding biological systems is challenging as the detail and complexity of such dynamic entities cannot be grasped using human ken and intuition. And disease states such as cancer further the difficulties in addressing living entities. Thus, investigators have created networks of cellular systems that encompass many components, connect those parts in some fashion, and then interrogate their usefulness for addressing predictability and/or to study the networks themselves (1). This has been well studied in the case of transcription factor networks that characterize the phenomenon of epithelial to mesenchymal transition (EMT) in cancer (2–6). Recent work demonstrated that the topology of transcription factor regulatory networks that were parameter free (using a Boolean approach) or were parameter agnostic (using a random parameter generator) was important in limiting changes in cell state that may promote disease progression (3).

While important findings can be gleaned from understanding regulation of gene expression, the effectors are not included in such modeling nor are site-specific details such as binding affinity or catalytic activities such as phosphorylation that are likely to influence a particular cell behavior. Therefore, the remainder of the Mini Review will focus on the modeling approaches that include the depth and likely parameters that may improve useful predictions.

Modeling cell behaviors using the mathematics underpinning biochemical reactions has been a research topic for decades when first approached by Tyson and others in understanding the cell cycle in the 1980s (7). It was and still is clear that cellular networks and pathways are dynamic and the overall contribution to cell behavior mattered. Since then, a vast array of biological models described by ordinary and partial differential equations (ODE, PDE) have been published, new software has been created to entice bench scientists to advance their findings computationally [for example, RuleBender (8) and Virtual Cell (9)], and now, an unprecedented amount of data to supply those models is available. For the oncology field, melding both mathematical models and omics data holds the promise to select the most effective molecular targets and any concurrent biomarkers. The future promise consists of whole cell models that lead to decisions of personalized therapies needed to predict tumor regression.

At the cellular level, mathematical models describe reaction networks (e.g., signal transduction pathways and metabolic cycles) effectively, typically with output that is difficult to predict. For example, binding of a growth factor to its receptor is a reaction that is dependent on characteristics such as binding affinity and concentration and leads to downstream events of interest occurring with time. Further, if positive and/or negative feedback loops are considered, cellular function becomes less predictable by inspection, if at all. But, operationally, in this example, the change in concentration of bound and unbound receptors and growth factor with respect to time can be tracked through species in the ODEs and downstream effects delineated. Furthermore, properties such as molecular diffusion can be described with PDEs, thus including spatial considerations as well. Many other biological properties have been described and include cytoskeleton formation (10), vesicular transport (11), and gene expression (12). Hundreds of such mathematical models have been deposited in the BioModels database (13) (RIDD: SCR_001993).

It is the premise of this Mini Review to describe how computational dynamic ODE models describing biochemical reaction networks can be analyzed by Machine Learning (ML) algorithms capable of predicting desirable outcomes for the two of the present challenges in oncology: discovering biomarkers associated with positive patient outcomes and novel molecular targets not generally considered druggable.

SELECTED ODE MODELS

As experimentalists have supplied an understanding of the essential knowledge of biochemical pathways underscoring cell behavior, ODE models offered predictions on some of the more

interesting aspects. For example, stimulus response, in general, was thought to occur in a linear fashion such that the greater concentration of a growth factor, the greater the response. However, some components in pathways appeared not to follow this linear rule. Rather, an all-or-none response occurred. ODE modeling supplied a mechanism for such ultrasensitivity, demonstrating a phenomenon that occurs when proteins with enzymatic function are acting at saturation (14, 15). Other important discoveries ensued and included uncovering bistability, a behavior dependent on initial conditions. Bistability is defined as having two stable states at one stimulus level (16, 17). Parenthetically, bistability has also been defined to describe whole cell states such as those found in EMT and is distinct from that which is noted here (3–5). Functionally, bistability at the biochemical reaction level is of interest because, dependent on conditions such as saturation of an enzyme and high initial stimulus, cellular response can resemble a toggle switch in the on position. Thus, immediate downregulation may not be needed in these instances (16, 18). Finally, oscillatory behavior of important transcription factors such p53 and their regulators have been demonstrated through ODE modeling and in cells (19, 20). In addition, oscillatory levels of phosphorylated kinases have been characterized by an amplitude and frequency and were shown to be regulated and define outcomes such as the decision to proliferate (21). Below are two relevant examples of ODE models that can be exploited to discover biomarkers and molecular targets in oncology.

MAPK Signaling

One of the most well-studied signal transduction pathways in oncology is that which triggers cell proliferation *via* growth factor stimulation. In one such reaction network, EGF, a growth factor, binds to its cognate receptor, which in turn, dimerizes and activates its kinase domain through conformational adjustments. Trans-autophosphorylation occurs next which forms the initial sites for adapter binding. GRB2 binding through its SH2 domain to phosphorylated tyrosine then binds SOS, a GTPase exchange factor. Critically, SOS replaces GDP with GTP on RAS, thus activating it for downstream binding of RAF, a kinase. A kinase cascade ensues ultimately leading to a phosphorylated kinase (phospho-ERK1/2) capable of triggering gene transcription necessary for cell growth. Importantly, several positive and negative feedback loops regulate the pathway and when considered in an ODE model, oscillatory behavior of phosphorylated ERK1/2 protein level occurs. While this description adequately corresponds to non-oncogenic signaling, disruption in the reaction rates by mutations leads to unpredictable outcomes dependent on pathway protein levels (22, 23). **Figure 1** (upper portion) depicts a limited Contact Map of MAPK signaling.

Cell Cycle Arrest or Apoptosis Decision

Cell decisions regulated by p53 have been readily modeled as it is of deep interest to determine the conditions in which the outcome of cell cycle arrest or apoptosis occurs following genomic insult. This is of particular importance since p53 is

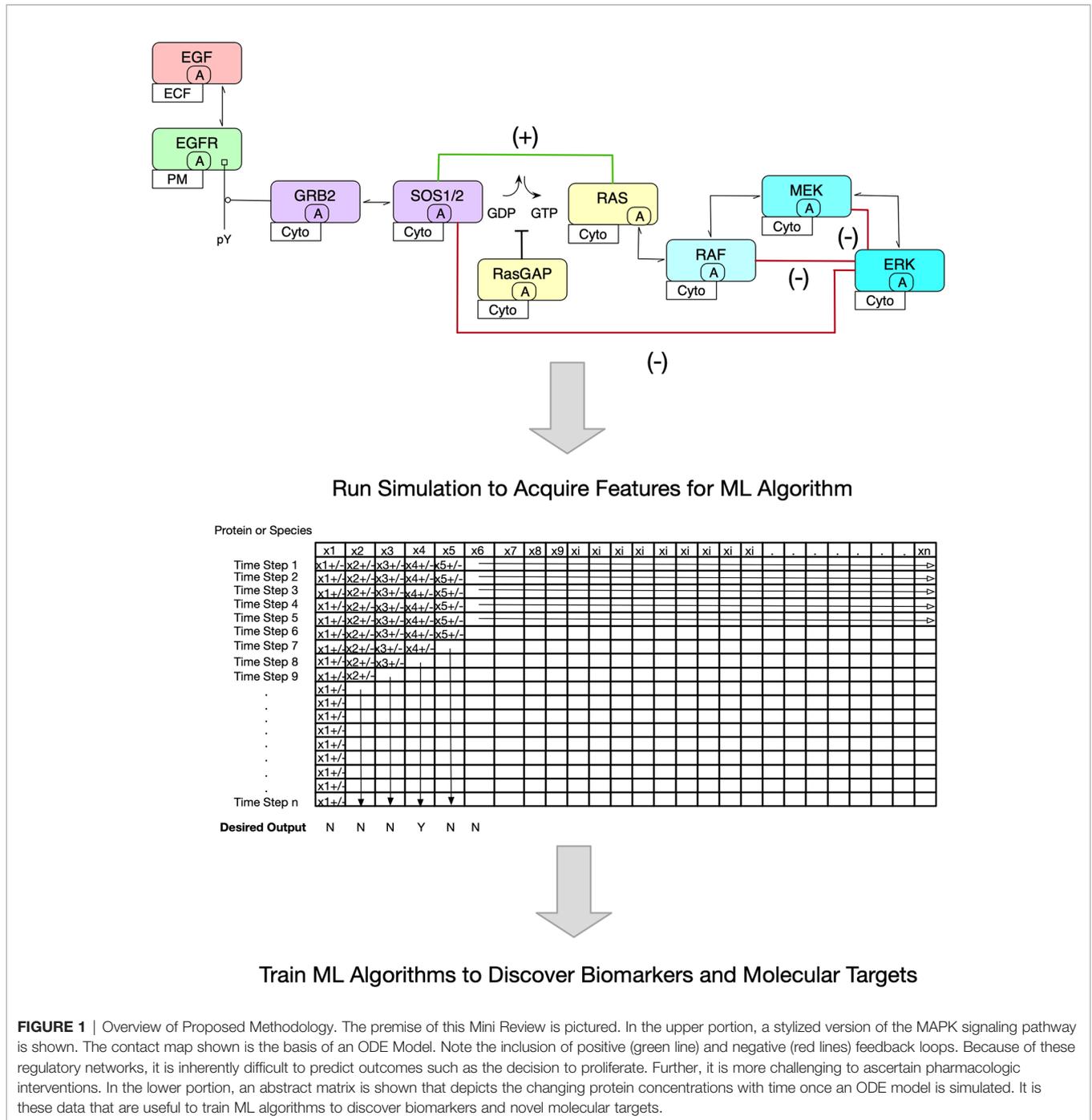


FIGURE 1 | Overview of Proposed Methodology. The premise of this Mini Review is pictured. In the upper portion, a stylized version of the MAPK signaling pathway is shown. The contact map shown is the basis of an ODE Model. Note the inclusion of positive (green line) and negative (red lines) feedback loops. Because of these regulatory networks, it is inherently difficult to predict outcomes such as the decision to proliferate. Further, it is more challenging to ascertain pharmacologic interventions. In the lower portion, an abstract matrix is shown that depicts the changing protein concentrations with time once an ODE model is simulated. It is these data that are useful to train ML algorithms to discover biomarkers and novel molecular targets.

commonly mutated in cancer and underpins tumorigenesis. Lipniacki's group developed a sophisticated model wherein irradiation induced DNA damage, triggering p53 transcription and subsequent downstream reactions (24, 25). The p53 network was created using the rule-based modeling technique that defines biochemical interactions such as protein complex formation and enzymatic activities. In the three-stage model, a steady state was achieved first through simulation and solutions through deterministic algorithms. Next, irradiation was modeled over a

range (1-10 Gy). In the last relaxation phase, cell behavior was characterized. Notably, the predicted outcomes were contingent on the degree of irradiation. Oscillatory behavior of p53 levels was observed as has been demonstrated *in vitro* and was found to define the conditions that lead to programmed cell death. Finally, their findings provided support for the heterogeneity in response to DNA damage observed in tumor cell lines. [N.B. For an overview of the methodology of ODE modeling in apoptosis, see (26)].

PROTEOMIC DATABASES

One of the challenges of ODE modeling is the need to insert quantitative parameters such as protein concentration in the reaction network. Parenthetically, precise measurements of reaction rates, enzymatic activity, and binding affinities are also critically important in computational modeling. The advent of critical technologies that can determine such include labeled and label-free assays such as SILAC and LC-MS/MS, respectively (27, 28). Both tumor cell lines and tumor tissue have been studied in this fashion and databases exist to exploit (29–32). And these quantitative proteomics efforts can evaluate post-translational modifications as well (33). The direct analysis of datasets from these studies has identified biomarkers, therapeutic targets, and drug resistance mechanisms. Further, and importantly, direct interaction networks can be constructed, but offer a static interpretation of a dynamic living cell.

The National Cancer Institute's Office of Cancer Clinical Proteomics developed one such database that contains proteomic evaluations from 13 different tumor types collected since 2006 (31, 32, 34) (RIDD: SCR_017135). A series of studies merged both proteomic and genomic data to better classify actionable mutations that are expressed as proteins with certainty rather than through sequence alone. This is even more critical since transcriptomics does not necessarily confirm translation in as high as 50% of all transcripts measured. But importantly, studies utilizing this database have provided important new classifications in cancer histologies. One study evaluated the proteogenomics of pediatric brain tumors and found new subgroups with wild type BRAF and novel networks that overlap with the mutant gene (35). Thus, new therapeutic trials could be proposed for these challenging tumor types.

SILAC methods have been melded with mathematical modeling in the study of dynamic systems. For example, Yilmaz et al. investigated proteosomal processing of NF κ B subunits in mouse embryonic fibroblasts *via* labeling studies and mathematical modeling to discover the dynamics of the system under activation (36). In another study, CHO cell extracts were processed in a glycoproteomic approach to understand N-glycan processing and found a kinetic description of the pathway (37). Finally, global protein synthesis rates were studied by pulse labeling and compared to mRNA synthesis rates in mathematical models, thus, providing critical quantitative parameters useful for future studies as well (38). In summary, SILAC and LC-MS/MS methods have supported mathematical modeling and hold further promise.

MACHINE LEARNING

Applying ML algorithms to uncover cancer diagnoses from histologies, to determine therapeutic decisions, and predict outcomes is becoming pervasive in light of omics studies (39, 40). ML, in an overview, can sort through vast inputs and discover connections not likely to be found by human inspection. An alternative application for ML in oncology could include the development of hypotheses subject to future study. Thus, ML is suited to intake vast inputs and to realize relationships not previously

expected. It is the thesis of this Mini Review to offer ways to meld ODE modeling and ML to discover biomarkers and molecular targets, ultimately aiding drug discovery and predictive oncology.

Biomarker and Molecular Target Discovery

It is instructive to describe a published ODE model in more detail for EGFR signaling encompassing the MAPK pathway (**Figure 1**) to demonstrate how it might be utilized in ML with existing proteomic databases. In two models published by Creamer and colleagues (41) and Kochanczyk and colleagues (22), the basic signaling pathway is described for EGF binding to the family of cognate receptor tyrosine kinases with subsequent dimerization triggering autophosphorylation on their intracellular tails at multiple sites. Next, adaptors bind with some affinity *via* SH2 or PTB domains. In canonical MAPK signaling, bound GRB2 anchors SOS1/2, a GTPase exchange protein. During activation, SOS1/2 binds RAS and exchanges GDP for GTP, and so activates RAS. A kinase cascade follows RAF1 dimerization and binding to RAS. RAF1 phosphorylates MEK1/2 and it, in turn, phosphorylates ERK1/2 which translocates to the nucleus to modulate transcription regulating essential genes for cell proliferation. The ODE model of EGFR signaling as described here includes reaction rates that underlie binding affinities, turnover rates, phosphorylation and dephosphorylation rates, and both positive and negative feedback loops (22). Including these important regulatory features (the loops) is critical since unexpected behaviors emerge such as oscillatory ERK1/2 phosphorylation levels and multiple states that include a steady state, a monostable one, or a bistable one dependent on the mathematical nature of the feedback loops. As an aside, even at this basic understanding, choosing a molecular target would be challenging.

Establishing the reaction network such as the one above by ODE equations can be readily accomplished with rule-based modeling using BioNetGen, a programming language and software that can further simulate the model over time (8). The simulations can be deterministically or stochastically solved with well accepted algorithms. Deterministic solutions are reproducible, resulting in the same output with every simulation. In contrast, stochastic simulations apply randomness to the solution, hence output is variable within a certain distribution and can be averaged. However, it is important to note that biological systems are noisy and the latter solutions may be more relevant. Simple output includes changes in species (molecule) abundance at each time step modeled. In the end, a vast amount of data is generated and fairly complex models with thousands of reactions can be coded (42). **Figure 1** (lower portion) pictorially describes a matrix of protein concentrations (in the columns) that change at each time step during a deterministic simulation (in the rows).

Parameterization of ODE models is challenging and typically, the scientific literature can provide many (43). It is anticipated that the proteomic databases described herein will readily provide protein abundance for ODE models and thus, may reflect both normal and disease states in separate models. Databases such as BRENDA (44, 45) (RIDD: SCR_002997) and Binding Database (46) contain curated reaction rates and affinities culled from the published literature. Thus, with a completed ODE model of interest, experimental protein abundance and measured parameters underlie their usefulness.

For ML algorithms, each time step can be represented by copy number (abundance) for an individual protein and thus, become features. What would be of interest for both biomarker and molecular target discovery is which state (i.e., time point and copy numbers) would be predictive of a desired outcome such as inhibition of cell proliferation and/or induction of cell death through one or more of the many known mechanisms. In order to achieve this, a training set would be needed that describes a signal transduction pathway or pathways anticipated to be central to cancer pathogenesis and clinically relevant. For example, the p53 ODE model such as the one described above intersects a critical cellular decision in light of DNA damage, determining cell cycle arrest or programmed cell death. The proteins in this training set model would be derived from tumor cell lines or tissues. Next a simulation would be completed resulting in a matrix of protein abundance at each time step (Figure 1). Now, the investigator has hundreds, if not thousands of models with and without the desired outcome that act as the training dataset for ML. The first analysis of such a ML model would be to find biomarkers (i.e., abundances of particular proteins) that correlate with the predicted outcome.

An alternative approach utilizing the same ODE model can aid in the discovery of novel molecular targets. In this case, proteins can be knocked out virtually (individually) through simple programming, a simulation run for each knock-out, and outcome collected. The ML algorithm would identify the connection between the presumptive molecular target and programmed cell death in this instance. Thus, a virtual high throughput screen has been completed using only computational effort.

CONCLUSION

It has been proposed in this Mini Review to apply ML algorithms to discover biomarkers and molecular targets through the

creation of ODE models of signaling pathways in cancer (47). While ODE modeling is more labor intensive than the ML analysis, the complex systems of cancer cell biology can be studied in this novel way leading to knowledge that will be readily apply to the pharmaceutical challenges ahead. In addition, ongoing advances in ODE modeling combined with tissue level simulations also show promise. For example, mathematical models of metabolism and cell proliferation indicated new molecular targets (48) and such multicellular models can further predict outcomes such as necrosis and growth arrest (49), cancer cell migration (50), and immune cell invasion (51). It can be envisioned that the computational efforts described herein can contribute to proposed Digital Twins for personalized medicine in cancer (52).

AUTHOR CONTRIBUTIONS

SM wrote the manuscript, completed the revisions, and approved of the submitted and revised version.

FUNDING

SM was funded by Mount St. Mary's University as Assistant Professor and BioSystems Strategies, LLC as Founder and CEO.

ACKNOWLEDGMENTS

The author wishes to acknowledge the expert advice from Dr. M. Kathy Jung in editing and revising the manuscript. In addition, the author thanks the generous support from Dr. Eric Stahlberg whose insights underpin the ideas set forth herein.

REFERENCES

- Kauffman S. Gene Regulation Networks: A Theory For Their Global Structure and Behaviors. *Curr Top Dev Biol* (1971) 6:145–82. doi: 10.1016/S0070-2153(08)60640-7
- Steinway SN, Zañudo JGT, Michel PJ, Feith DJ, Loughran TP, Albert R. Combinatorial Interventions Inhibit Tgf β -Driven Epithelial-to-Mesenchymal Transition and Support Hybrid Cellular Phenotypes. *NPJ Syst Biol Appl* (2015) 1:15014. doi: 10.1038/npjbsa.2015.14
- Hari K, Sabuwala B, Subramani BV, La Porta CAM, Zapperi S, Font-Clos F, et al. Identifying Inhibitors of Epithelial-Mesenchymal Plasticity Using a Network Topology-Based Approach. *NPJ Syst Biol Appl* (2020) 6(1):15. doi: 10.1038/s41540-020-0132-1
- Pillai M, Jolly MK. Systems-Level Network Modeling Deciphers the Master Regulators of Phenotypic Plasticity and Heterogeneity in Melanoma. *iScience* (2021) 24(10):103111. doi: 10.1016/j.isci.2021.103111
- Udyavar AR, Wooten DJ, Hoeksema M, Bansal M, Califano A, Estrada L, et al. Novel Hybrid Phenotype Revealed in Small Cell Lung Cancer by a Transcription Factor Network Model That Can Explain Tumor Heterogeneity. *Cancer Res* (2017) 77(5):1063–74. doi: 10.1158/0008-5472.CAN-16-1467
- Lang J, Nie Q, Li C. Landscape and Kinetic Path Quantify Critical Transitions in Epithelial-Mesenchymal Transition. *Biophys J* (2021) 120(20):4484–500. doi: 10.1016/j.bpj.2021.08.043
- Tyson JJ, Hannsgen KB. Cell Growth and Division: A Deterministic/Probabilistic Model of the Cell Cycle. *J Math Biol* (1986) 23(2):231–46. doi: 10.1007/BF00276959
- Blinov ML, Faeder JR, Goldstein B, Hlavacek WS. BioNetGen: Software for Rule-Based Modeling of Signal Transduction Based on the Interactions of Molecular Domains. *Bioinformatics* (2004) 20(17):3289–91. doi: 10.1093/bioinformatics/bth378
- Moraru II, Schaff JC, Slepchenko BM, Blinov ML, Morgan F, Lakshminarayana A, et al. Virtual Cell Modelling and Simulation Software Environment. *IET Syst Biol* (2008) 2(5):352–62. doi: 10.1049/iet-syb:20080102
- Novak IL, Slepchenko BM, Mogilner A, Loew LM. Cooperativity Between Cell Contractility and Adhesion. *Phys Rev Lett* (2004) 93(26 Pt 1):268109. doi: 10.1103/PhysRevLett.93.268109
- Shea SM, Karnovsky MJ, Bossert WH. Vesicular Transport Across Endothelium: Simulation of a Diffusion Model. *J Theor Biol* (1969) 24(1):30–42. doi: 10.1016/S0022-5193(69)80004-4
- Larson DR, Singer RH, Zenklusen D. A Single Molecule View of Gene Expression. *Trends Cell Biol* (2009) 19(11):630–7. doi: 10.1016/j.tcb.2009.08.008
- Malik-Sheriff RS, Glont M, Nguyen TVN, Tiwari K, Roberts MG, Xavier A, et al. BioModels-15 Years of Sharing Computational Models in Life Science. *Nucleic Acids Res* (2020) 48(D1):D407–15. doi: 10.1093/nar/gkz1055
- Goldbeter A, Koshland DE. An Amplified Sensitivity Arising From Covalent Modification in Biological Systems. *Proc Natl Acad Sci USA* (1981) 78(11):6840–4. doi: 10.1073/pnas.78.11.6840

15. Kim SY, Ferrell JE. Substrate Competition as a Source of Ultrasensitivity in the Inactivation of Wee1. *Cell* (2007) 128(6):1133–45. doi: 10.1016/j.cell.2007.01.039
16. Pomerening JR, Sontag ED, Ferrell JE. Building a Cell Cycle Oscillator: Hysteresis and Bistability in the Activation of Cdc2. *Nat Cell Biol* (2003) 5(4):346–51. doi: 10.1038/ncb954
17. Ferrell JE, Ha SH. Ultrasensitivity Part III: Cascades, Bistable Switches, and Oscillators. *Trends Biochem Sci* (2014) 39(12):612–8. doi: 10.1016/j.tibs.2014.10.002
18. Trunnell NB, Poon AC, Kim SY, Ferrell JE. Ultrasensitivity in the Regulation of Cdc25C by Cdk1. *Mol Cell* (2011) 41(3):263–74. doi: 10.1016/j.molcel.2011.01.012
19. Ma L, Wagner J, Rice JJ, Hu W, Levine AJ, Stolovitzky GA. A Plausible Model for the Digital Response of P53 to DNA Damage. *Proc Natl Acad Sci USA* (2005) 102(40):14266–71. doi: 10.1073/pnas.0501352102
20. Hat B, Kochańczyk M, Bogdał MN, Lipniacki T. Feedbacks, Bifurcations, and Cell Fate Decision-Making in the P53 System. *PLoS Comput Biol* (2016) 12(2):e1004787. doi: 10.1371/journal.pcbi.1004787
21. Albeck JG, Mills GB, Brugge JS. Frequency-Modulated Pulses of ERK Activity Transmit Quantitative Proliferation Signals. *Mol Cell* (2013) 49(2):249–61. doi: 10.1016/j.molcel.2012.11.002
22. Kochańczyk M, Kocieńiewski P, Kozłowska E, Jaruszewicz-Błońska J, Sparta B, Pargett M, et al. Relaxation Oscillations and Hierarchy of Feedbacks in MAPK Signaling. *Sci Rep* (2017) 7:38244–59. doi: 10.1038/srep38244
23. Stites EC, Shaw AS. Quantitative Systems Pharmacology Analysis of KRAS G12C Covalent Inhibitors. *CPT Pharmacometrics Syst Pharmacol* (2018) 7(5):342–51. doi: 10.1002/psp4.12291
24. Bogdał MN, Hat B, Kochańczyk M, Lipniacki T. Levels of Pro-Apoptotic Regulator Bad and Anti-Apoptotic Regulator Bcl-xL Determine the Type of the Apoptotic Logic Gate. *BMC Syst Biol* (2013) 7:67. doi: 10.1186/1752-0509-7-67
25. Hat B, Jaruszewicz-Błońska J, Lipniacki T. Model-Based Optimization of Combination Protocols for Irradiation-Insensitive Cancers. *Sci Rep* (2020) 10(1):12652. doi: 10.1038/s41598-020-69380-6
26. Rehm M, Prehn JH. Systems Modelling Methodology for the Analysis of Apoptosis Signal Transduction and Cell Death Decisions. *Methods* (2013) 61(2):165–73. doi: 10.1016/j.jymeth.2013.04.007
27. Li N, Li J, Desiderio DM, Zhan X. SILAC Quantitative Proteomics Analysis of Ivermectin-Related Proteomic Profiling and Molecular Network Alterations in Human Ovarian Cancer Cells. *J Mass Spectrom* (2021) 56(1):e4659. doi: 10.1002/jms.4659
28. Su F, Zhou F-F, Zhang T, Wang D-W, Zhao D, Hou X-M, et al. Quantitative Proteomics Identified 3 Oxidative Phosphorylation Genes With Clinical Prognostic Significance in Gastric Cancer. *J Cell Mol Med* (2020) 24(18):10842–54. doi: 10.1111/jcmm.15712
29. Ellis MJ, Gillette M, Carr SA, Paulovich AG, Smith RD, Rodland KK, et al. Connecting Genomic Alterations to Cancer Biology With Proteomics: The NCI Clinical Proteomic Tumor Analysis Consortium. *Cancer Discov* (2013) 3(10):1108–12. doi: 10.1158/2159-8290.CD-13-0219
30. Edwards NJ, Oberti M, Thangudu RR, Cai S, McGarvey PB, Jacob S, et al. The CPTAC Data Portal: A Resource for Cancer Proteomics Research. *J Proteome Res* (2015) 14(6):2707–13. doi: 10.1021/pr501254j
31. Guo T, Luna A, Rajapakse VN, Koh CC, Wu Z, Liu W, et al. Quantitative Proteome Landscape of the NCI-60 Cancer Cell Lines. *iScience* (2019) 21:664–80. doi: 10.1016/j.isci.2019.10.059
32. Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, et al. Next-Generation Characterization of the Cancer Cell Line Encyclopedia. *Nature* (2019) 569(7757):503–8. doi: 10.1038/s41586-019-1186-3
33. Jiang Y, Sun A, Zhao Y, Ying W, Sun H, Yang X, et al. Proteomics Identifies New Therapeutic Targets of Early-Stage Hepatocellular Carcinoma. *Nature* (2019) 567(7747):257–61. doi: 10.1038/s41586-019-0987-8
34. Rodriguez H, Zenklusen JC, Staudt LM, Doroshow JH, Lowy DR. The Next Horizon in Precision Oncology: Proteogenomics to Inform Cancer Diagnosis and Treatment. *Cell* (2021) 184(7):1661–70. doi: 10.1016/j.cell.2021.02.055
35. Petralia F, Tignor N, Reva B, Kopytyra M, Chowdhury S, Rykunov D, et al. Integrated Proteogenomic Characterization Across Major Histological Types of Pediatric Brain Cancer. *Cell* (2020) 183(7):1962–85.e31. doi: 10.1016/j.cell.2020.10.044
36. Yilmaz ZB, Kofahl B, Beaudette P, Baum K, Ipenberg I, Weih F, et al. Quantitative Dissection and Modeling of the NF- κ B P100-P105 Module Reveals Interdependent Precursor Proteolysis. *Cell Rep* (2014) 9(5):1756–69. doi: 10.1016/j.celrep.2014.11.014
37. Arigoni-Affolter I, Scibona E, Lin C-W, Brühlmann D, Souquet J, Broly H, et al. Mechanistic Reconstruction of Glycoprotein Secretion Through Monitoring of Intracellular N-Glycan Processing. *Sci Adv* (2019) 5(11):eaax8930. doi: 10.1126/sciadv.aax8930
38. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global Quantification of Mammalian Gene Expression Control. *Nature* (2011) 473(7347):337–42. doi: 10.1038/nature10098
39. Schneider P, Walters WP, Plowright AT, Sieroka N, Listgarten J, Goodnow RA, et al. Rethinking Drug Design in the Artificial Intelligence Era. *Nat Rev Drug Discov* (2020) 19(5):353–64. doi: 10.1038/s41573-019-0050-3
40. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine Learning in Medicine: A Practical Introduction. *BMC Med Res Methodol* (2019) 19(1):64. doi: 10.1186/s12874-019-0681-4
41. Creamer MS, Stites EC, Aziz M, Cahill JA, Tan CW, Berens ME, et al. Specification, Annotation, Visualization and Simulation of a Large Rule-Based Model for ERBB Receptor Signaling. *BMC Syst Biol* (2012) 6:107. doi: 10.1186/1752-0509-6-107
42. Dolan DWP, Zupanic A, Nelson G, Hall P, Miwa S, Kirkwood TBL, et al. Integrated Stochastic Model of DNA Damage Repair by Non-Homologous End Joining and P53/P21-Mediated Early Senescence Signalling. *PLoS Comput Biol* (2015) 11(5):e1004246. doi: 10.1371/journal.pcbi.1004246
43. Mitra ED, Hlavacek WS. Parameter Estimation and Uncertainty Quantification for Systems Biology Models. *Curr Opin Syst Biol* (2019) 18:9–18. doi: 10.1016/j.coisb.2019.10.006
44. Schomburg I, Chang A, Schomburg D. BRENDA, Enzyme Data and Metabolic Information. *Nucleic Acids Res* (2002) 30(1):47–9. doi: 10.1093/nar/30.1.47
45. Chang A, Jeske L, Ulbrich S, Hofmann J, Kobltz J, Schomburg I, et al. BRENDA, the ELIXIR Core Data Resource in 2021: New Developments and Updates. *Nucleic Acids Res* (2021) 49(D1):D498–508. doi: 10.1093/nar/gkaa1025
46. Gilson MK, Liu T, Baitaluk M, Nicola G, Hwang L, Chong J. BindingDB in 2015: A Public Database for Medicinal Chemistry, Computational Chemistry and Systems Pharmacology. *Nucleic Acids Res* (2016) 44(D1):D1045–53. doi: 10.1093/nar/gkv1072
47. Patterson EA, Whelan MP. A Framework to Establish Credibility of Computational Models in Biology. *Prog Biophys Mol Biol* (2017) 129:13–9. doi: 10.1016/j.pbiomolbio.2016.08.007
48. Roy M, Finley SD. Computational Model Predicts the Effects of Targeting Cellular Metabolism in Pancreatic Cancer. *Front Physiol* (2017) 8:217. doi: 10.3389/fphys.2017.00217
49. Ghaffarizadeh A, Heiland R, Friedman SH, Mumenthaler SM, Macklin P. PhysiCell: An Open Source Physics-Based Cell Simulator for 3-D Multicellular Systems. *PLoS Comput Biol* (2018) 14(2):e1005991. doi: 10.1371/journal.pcbi.1005991
50. Schumacher LJ, Maini PK, Baker RE. Semblance of Heterogeneity in Collective Cell Migration. *Cell Syst* (2017) 5(2):119–27. doi: 10.1016/j.cels.2017.06.006
51. Vipond O, Bull JA, Macklin PS, Tillmann U, Pugh CW, Byrne HM, et al. Multiparameter Persistent Homology Landscapes Identify Immune Cell Spatial Patterns in Tumors. *Proc Natl Acad Sci USA* (2021) 118(41):e2102166118. doi: 10.1073/pnas.2102166118
52. Hernandez-Boussard T, Macklin P, Greenspan EJ, Gryshuk AL, Stahlberg E, Syeda-Mahmood T, et al. Digital Twins for Predictive Oncology Will be a Paradigm Shift for Precision Cancer Care. *Nat Med* (2021) 27(12):2065–6. doi: 10.1038/s41591-021-01558-5

Conflict of Interest: The author had a position as Guest Researcher at Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, LLC, and is Founder and CEO of BioSystems Strategies, LLC.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Mertins. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.